*Language and Cognition* (2024), 1–22
doi:10.1017/langcog.2024.44

ARTICLE

# Learning novel words for motion by speakers of structurally different languages

Irmak Su Tütüncü[1], Samantha N. Emerson[2], Jing Paul[3], Murat Şengül[4] and Şeyda Özçalışkan[1]

[1]Psychology Department, Georgia State University, Atlanta, GA, USA; [2]Training, Learning & Readiness Division, Aptima, Inc., Woburn, MA, USA; [3]Asian Studies, Agnes Scott College, Decatur, GA, USA and [4]Department of Turkish Language Education, Nevşehir Hacı Bektaş Veli University, Nevşehir, Türkiye
**Corresponding author**: Irmak Su Tütüncü; Email: ittnc1@gsu.edu

(Received 05 January 2024; Revised 23 April 2024; Accepted 30 May 2024)

## Abstract
Speakers of different languages follow a three-way split in how they express motion events in speech—with a greater emphasis on manner in satellite-framed languages (English), path in verb-framed languages (Turkish), and comparable expression of manner and path in equipollently-framed languages (Chinese). According to the thinking-for-speaking account, these language-specific patterns can affect speakers' representation of motion events but only when verbalizing the event. In this study, we asked whether language might influence learning novel words, particularly when the words were accompanied with gestures. We examined effects of language type (equipollent-framed: Chinese, satellite-framed: English, verb-framed: Turkish) and modality (speech-only, gesture+speech) on learning pseudowords for motion (manner, path). Our results showed that speakers of all three languages learned pseudowords for manner and path but with lower accuracy scores and slower rates of learning by Chinese speakers. Regardless of the language they spoke, participants learned manner words more accurately than path words, but with no added benefits of instruction with gesture+speech over speech-only. Taken together, our study extends the lack of language effect on nonverbal representation of events when not speaking to the domain of novel word learning across structurally different languages.

**Keywords:** cross-linguistic difference; gesture; motion events; multi-modal word learning

## 1. Introduction

The expression of motion events differs systematically across the world's languages, with pronounced differences in the expression of manner (i.e., how one moves) and path (i.e., the direction one moves) components of motion (Talmy, 2000). Adult speakers show strong cross-linguistic differences, following the language-specific patterns in their speech about motion (Slobin, 2004). The language-specific patterns also influence the way speakers think about motion events, particularly when they are verbalizing the event. The online effect of language on nonverbal representation of events during verbalization, originally proposed by Slobin (1996) as the 'thinking for speaking account', has been shown across several studies that used a variety of nonverbal measures, including co-speech gesture (e.g., Özçalışkan et al., 2016a, 2016b), categorization (e.g., Gennari et al., 2002), memory (e.g., Oh, 2003), and attention (e.g., Emerson et al., 2020). There is however research that suggests that the effect of language on thinking does not go beyond verbalization of the event, with no effect of language on nonverbal representations of events when speakers are tested with measures that do not require language-specific description of the event (e.g., Athanasopoulos & Bylund, 2013; Cardini, 2010; Özçalışkan et al., 2016a, 2018; Papafragou et al., 2002; Tütüncü et al., 2023).

In this study, we extended earlier work to the domain of *novel word learning*, asking whether the habitual patterns of motion expression in one's native language affect learning of verbal labels for motion events, particularly when those labels are presented with gesture. More specifically, using a word-learning paradigm that included training with or without gestures, we tested whether speakers of three structurally different languages (equipollently-framed: Chinese, satellite-framed: English, verb-framed: Turkish) would show language-specific effects when learning pseudowords that encoded manner or path of motion—without verbalization of the event in their native language—and whether this effect would become evident in both explicit (i.e., accuracy of behavioral response) and implicit (i.e., speed of behavioral response) measures for learning. If language influences nonverbal representation of events only during verbalization in one's native language—as suggested in Slobin's (1996) thinking for speaking account—then we would expect speakers of all three languages to learn pseudowords encoding manner and path equally well in a novel word-learning context as they are not constrained by the habitual patterns of motion expression in their native language during learning. If, on the other hand, language's effect on nonverbal representation goes beyond online production of native speech, then we would predict that speakers would differ in learning pseudowords for manner and path, showing an advantage in learning labels more frequently expressed in their native language—particularly when the labels are accompanied with gesture.

### 1.1. Cross-linguistic variability in talking about motion events

Spatial motion constitutes a core human experience; however, its expression shows strong cross-linguistic variability. As originally proposed by Talmy (1985, 2000) and later expanded by Slobin (2004), the world's languages can be categorized into different types along a tertiary split between satellite-framed (S-language; e.g., English, Polish), verb-framed (V-language; e.g., Turkish, Spanish) and equipollently-framed (E-language; e.g., Chinese, Thai) languages based primarily on the expression of path of motion, which in turn has consequences for the expression of manner of motion. In S-languages, path is expressed in a particle outside the verb, reserving the main verb for conveying manner as in: *he runs* (manner) *into* (path) *the house.* In contrast, in V-languages the verb encodes path of motion, and manner is optionally expressed outside the verb in a secondary lexical element as in: *ev-e girer koşarak* = he enters (path) house-to by running (manner). The serial-verb construction in E-languages, on the other hand, allows for expression of both manner and path information in the verb as in: *tā paojin fángzi* = He run (manner)-enter (path) house.

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

The preference for using the main verb for expressing path or manner of motion in these different languages has important consequences, particularly for the amount and diversity of manner and path verb production. As shown in earlier work, when describing motion events, adult speakers of S-languages (e.g., English, German, Dutch, Polish) use greater amounts and variety of manner verbs than adult V-language speakers (e.g., Basque, French, Spanish, Turkish)—a pattern that is reversed for the production of path verbs (e.g., Cardini, 2010; De Knop & Dirven, 2008; De Knop & Gallez, 2011; Hickmann et al., 2009; Ibarretxe-Antunano, 2009, 2012; Lewandowski & Özçalışkan, 2018, 2023; Naigles & Terrazas, 1998; Özçalışkan & Slobin, 1999, 2003; Tusun & Hendriks, 2019). In fact, V-language speakers frequently leave out manner information altogether from their motion descriptions by not expressing it either in the verb or outside the verb and primarily convey path of motion (e.g., Emerson et al., 2021; Özçalışkan, 2015). The relatively limited work on E-languages suggests that E-language speakers express manner and path information at comparable rates, as they have the option to express both in a serial verb construction (Chen & Guo, 2009; Paul, Emerson & Özçalışkan, 2022)—a pattern that contrasts with both S- and V-languages.

In summary, existing research on speech about motion suggests that adult speakers show strong but systematic cross-linguistic variation in their verbal expression of manner and path of motion, with greater encoding of manner in S-languages, path in V-languages, and comparable expression of manner and path in E-languages.

### 1.2. Cross-linguistic variability in thinking about motion events

The cross-linguistic variability evident in talking about motion events raises the possibility that speakers of these three types of languages might also think about motion in different ways. The existing research indeed suggests that language-specific pattern of motion expression has an effect on nonverbal representation of events; but this effect is evident when nonverbal tasks are accompanied by verbalization of the event in native language and *not* observable when the cognitive tasks are completed without verbalization (i.e., without speaking, hearing, or writing in one's native language). For example, an earlier study (Gennari et al., 2002), using a similarity judgment task, examined whether speakers of English (S-language) or Spanish (V-language) would show biases consistent with their language (manner for English, path for Spanish) in drawing similarities between events that depicted manner or path variations. The participants were presented with an initial event, followed by two test follow-up events: one showing a different path with the same manner (i.e., same-manner alternative) and the other showing a different manner with the same path (i.e., same-path alternative); they were then asked to pick the test event most similar to the original event, but after they described the original event in their native language. Spanish speakers were more likely than English speakers to choose same-path events as being more similar to the initial event—a pattern that was reversed for English speakers, thus suggesting an online effect of language on representation of events (as evidenced by similarity ratings). In another study, Oh (2003) examined whether speakers of English or Korean (V-language) differed in their memory for manner versus path components of motion events, using dynamic motion scenes in which the speakers were asked to verbalize the event in their language. English speakers not only expressed manner more frequently and in greater detail than Korean speakers, but also

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

showed better memory for subtle differences in manner of motion in the events that they described as compared to Korean adults. The effect of language on thinking has also been shown at the neural level in an event-related potential (ERP) task that involved reading motion descriptions in English or Spanish with manner and path verbs (Emerson et al., 2020). The speakers in the two languages showed different neural responses (P600 effects) when reading motion verbs for manner versus path, indicating that English speakers showed a greater expectancy for motion verbs to express manner while Spanish speakers showed a greater expectancy for motion verbs to express path—a neural pattern consistent with the language-specific patterns of motion expression in the two languages.

The differential bias to manner versus path in nonverbal tasks dissipates, however, when the experimental task does *not* involve verbalization of the event. An earlier study (Papafragou et al., 2008) examined eye gaze patterns of English (S-language) and Greek (V-language) speakers while they viewed motion event animations with manner and path components (e.g., a man skiing to a rocket) with or without verbalization of the event. The eye movements of the speakers focused more on scene components that are frequently encoded in their native language (manner in English, path in Greek) when verbalizing the event. However, the cross-linguistic differences were not evident when the speakers viewed the events without verbalization; they instead showed comparable gaze patterns to manner and path components of the same events, suggesting a lack of language effects when not verbalizing in their native language. In a similar vein, a recent study (Skordos et al., 2020) examined differences in memory for manner and path of motion among English and Greek speakers. The participants were asked to view short-animated motion clips (e.g., an alien driving a car towards a rock) quietly, without verbalization, and then remember the clips as best as they could for a later memory task. The results showed no effect of language on memory, with both Greek and English speakers remembering path of motion better than manner of motion. The lack of language effects was also evident in a study by Cardini (2010) with English and Italian (V-language) speakers. Participants in this study watched a target video of real people performing a motion event with manner and path (e.g., man climbs down a slide); they were then asked to judge the similarity of this original video to a test video that either matched the manner (e.g., man climbs up a slide) or the path (e.g., man slides down a slide) depicted in the original video but without verbalization of the event. The participants showed no effect of language in their similarity judgments, with similar performance in their matching responses for manner or path.

Another set of studies, using gesture production as a nonverbal measure, examined whether speakers of different languages would follow language-specific patterns in their gestures when producing gestures with verbalization (i.e., gesturing while speaking). Most of this earlier work showed that gestures mirror the patterns found in speech, thus showing an effect of language on nonverbal representation of events in gesture (Goldin-Meadow et al., 2008; Kita & Özyürek, 2003; Özçalışkan et al., 2016a). More specifically, speakers of S-languages combine manner and path of motion into the same gesture (e.g. wiggle fingers while moving from left to right to convey running along a given path) while V-language speakers predominantly express only path of motion in their gestures about motion (e.g., trace a line forward with finger to convey forward trajectory; Gullberg et al., 2008; Özçalışkan et al., 2016b, 2018). A more recent study (Tütüncü et al., 2023) extended these patterns to E-languages, comparing Chinese speakers to English and Turkish speakers in an animated motion event description task. Chinese speakers, who expressed manner and path together in the verb using serial verb constructions, also synthesized manner and path into a single gesture at rates greater than both English and Turkish speakers, thus extending the effect of language on co-speech gesture to E-languages.

Importantly, the effect of language was not evident when gestures were produced without verbalization (i.e., gesturing without speaking). Adult speakers of S- versus V-languages (i.e., English versus Turkish) gestured in the same way when describing event scenes solely with their hands without any verbalization of the event (Özçalişkan, 2016; Özçalışkan et al., 2016a, 2018). They all expressed manner and path of motion together in a single gesture (e.g., run fingers forward to convey running towards house) and at roughly comparable rates. This pattern was later extended to E-languages (i.e., Chinese), using an animated motion event description task without verbalization of the event (Tütüncü et al., 2023). Chinese speakers also expressed both manner and path in a single gesture, showing a pattern akin to the English and Turkish speakers in the study.

In summary, the existing work suggests an effect of language on thinking during verbalization of a motion event but *no effect* of language when *not verbalizing* the event in one's native language. Speakers show a bias consistent with their native language across a variety of nonverbal tasks—from co-speech gesture to memory—but only when asked to verbalize the motion event in their native language. This language-specific bias disappears when speakers complete the same nonverbal tasks without accompanying native speech, suggesting limits on the effects of language on the nonverbal representation of events.

### 1.3. Cross-linguistic variability in learning novel words about motion events

We know from earlier work that children learn language-specific patterns at an early age when exposed to their native language at birth. More specifically, children learning S-languages produce a greater number and variety of manner verbs, while children learning V-languages use a greater amount and variety of path verbs beginning around age 3-4 (Allen et al., 2007; Hickmann et al., 2009; Özçalışkan, 2009; Özçalışkan et al., 2024a; Özçalışkan & Slobin, 1999; Skordos & Papafragou, 2014; Smyder & Harrigan, 2021), thus suggesting early attunement to language-specific patterns in children's speech about motion events in native production contexts.

We know relatively less about the effect of language when learning novel words for motion as an adult. A few studies examined learning novel words for motion events embedded within sentences in speakers' native languages (e.g., *She is kradding; Ella está mecando = She is mec-ing*), thus allowing the verbalization of the events in speakers' native language. These studies found that S-language speakers were more likely to interpret a pseudoword as expressing manner and V-language speakers were more likely to interpret the same pseudoword as expressing path (e.g., English versus Spanish; Naigles & Terrazas, 1998; Greek versus English; Papafragou & Selimis, 2009; English versus Spanish versus Japanese; Maguire et al., 2010; English: Shafto et al., 2014). These findings thus suggest notable language-specific biases in assigning meaning to novel words during learning, but only when verbalizing the event in one's native language. This is consistent with the thinking for speaking account (Slobin, 1996).

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

There is only one cross-linguistic study to our knowledge that examined novel motion word learning using pseudowords without any accompanying verbalization. In this study, Kersten et al. (2010) asked English (S-language) and Spanish (V-language) speakers to categorize motion events animated with bug-like creatures that differed from each other in both their motion type (i.e., key categorical variable: path versus manner) and their appearance (i.e., distractor variable, e.g., color, body shape, number of legs), using either pseudowords or numbers as labels. Kersten et al. (2010) found that English speakers were more accurate than Spanish speakers in identifying the categorical pseudoword for a creature when the relevant feature was manner but not path of motion (regardless of label type), thus suggesting an effect of language on novel word learning that goes beyond the verbalization of the event. However, a later study (Emerson et al., 2016) that examined novel word learning by adult English speakers (S-language) without verbalization did not show better performance on learning words that encoded manner variations than path variations, suggesting a lack of language-specific effects in learning novel verbs that goes beyond verbalization. The pattern of findings for the possible effect of language on learning novel words that extends beyond verbalization of the event thus remains largely inconclusive.

Some of the previous work on novel word learning also examined whether speakers would benefit from gesture instruction when learning novel words. Earlier studies that primarily focused on cognitive tasks (e.g., mathematical equivalence problems, identifying symmetry, Piagetian conservation problems) showed that gesture could facilitate learning, especially if the task at hand is cognitively challenging for the learner (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Perry & Elder, 1997; Ping & Goldin-Meadow, 2008; Valenzeno et al., 2003). The better learning with gesture was attributed to the more direct and precise communication (Holle & Gunter, 2007; Kang & Tversky, 2016), reduced cognitive load (Goldin-Meadow & Alibali, 2013), and improved memory (Mathias et al., 2021) afforded by the accompaniment of gesture.

Only a few studies, however, examined the added benefit of instruction with gestures in learning novel words among adult speakers. Three such studies, taught adult English speakers new words in a language that they had no knowledge of (i.e., Chinese: Huang et al., 2019; Hungarian: Morett, 2014; Sweller et al., 2020) by using either speech-only or gesture with speech (i.e., gesture+speech) instruction. All three studies found that gesture aided novel word learning, with English speakers showing better learning of the novel words in another language when instructed with gesture+speech than with speech-only. A similar beneficial effect of gesture was also found in another study in which English speakers learned pseudowords embedded within English sentences (e.g., I got the *zek* from the library; Everyone should *fim* breakfast), with different types of gesture (e.g., iconic, beat) or without gesture and with different instruction types (simple versus complex instruction; Hupp & Gingras, 2016). Participants learned pseudowords taught with iconic gestures better than the ones taught without gesture (or with other gesture types) independent of the complexity of the instruction. This pattern was also evident in an fMRI study (Macedonia et al., 2011) where native German speakers learned Italian words that they have never encountered before better when instructed with meaningful gestures (i.e., iconic gestures that conveyed relevant semantic information) than with meaningless gestures (i.e., gestures that are not related to the word's semantics). The brain activity recorded in the two conditions was different: meaningful gestures activated premotor cortices; while meaningless gestures elicited a network associated with cognitive control. These findings thus suggest that memory for newly learned words is mostly driven by the motor image that matches an underlying representation of the word's semantics rather than a mere effect of motor activity. At the same time, there is also research that suggests that even meaningless beat gestures can aid novel word learning among second language learners. In an earlier study (Kushch et al., 2018), Catalan dominant native speakers learned novel words in Russian with prosodic prominence better when the words were accompanied with beat gestures emphasizing the pronunciation of the words.

Some of the other work, on the other hand, suggested no effect of gesture on learning novel words. For instance, Emerson et al. (2016) taught pseudowords to adult English speakers, either with speech or with gesture+speech, and showed no modality-based differences in learning. In another study, Kelly and Lee (2012) taught adult English speakers novel words in a language that they had no knowledge of (i.e., Japanese) with or without iconic gesture instruction (e.g., gesturing 'stay' while saying *ite* = stay in Japanese versus saying *ite* = stay in Japanese with no accompanying gesture). Some of the words constituted phonologically easy pairs (e.g., "tate-butta" = stand-hit) while others were more difficult ("*ite* – *itte*" = stay – go); the results showed a beneficial effect of instruction with gesture but only for word pairs that were phonologically easier, suggesting that task difficulty might be an additional factor in determining gesture's role in word learning. It is also important to note here that, the beneficial effect of gesture in learning novel words was more evident in studies that also taught novel labels for objects and features (e.g., Kushch et al., 2018; Morett, 2014) but was less evident in studies that aimed to teach novel labels for motion (e.g., Emerson et al., 2016; Kelly & Lee, 2012).

In summary, research on novel word learning in the context of motion events—with or without gesture—remains sparse, with a few studies suggesting an effect of language on learning novel words for motion, particularly when the novel words were embedded within an event description (i.e., verbalized) in the speakers' native language. The beneficial effect of gestures on word learning, on the other hand, remains unclear, with limited and largely inconclusive results, almost all based on native English speakers. This, in turn, highlights the need for future studies that examine gesture's effect on learning novel words across speakers of a greater variety of languages.

### 1.4. Current study

Speakers of different languages vary in the way they talk about motion events, showing a three-way split in the expression of the manner and path components of motion (Slobin, 2004; Talmy, 2000). These differences have been shown to affect nonverbal representation of motion events but only during verbalization of the event in one's native language—a pattern consistent with the 'thinking for speaking account' (Slobin, 1996). At the same time, studies that focus on word learning in structurally different languages—with or without gesture—remain sparse, with no research to date examining the effect of modality and language type on motion-word learning in a single research design. In this study, we used a comprehensive framework to understand the factors that contribute to variability in learning novel words for motion in adult speakers in three structurally different languages. More specifically, we examined whether learning novel words for manner or path is affected by

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

language (Chinese, English, Turkish) or modality (speech-only, gesture+speech) in a learning task that does not involve verbalization of the event. We asked two main questions:

(1) We first asked whether speakers of the three languages would show an effect of *motion type* in their learning of words for motion. We had a two-way prediction: If language has an effect on nonverbal representation of events only during verbalization in native language—as suggested by the 'thinking for speaking account'—then we would predict that Chinese, English and Turkish speakers would *not* differ in their learning of novel words for manner or path (as they are not using their native language to verbalize the motion events during the learning task). However, if language's effect on nonverbal representation goes beyond verbalization in native language, then we would predict that speakers of the three languages would differ in their learning of the pseudowords, with better learning of words for manner in English, path in Turkish, and similar levels of learning for manner and path in Chinese.

(2) We next asked whether speakers of the three languages would show an effect of *modality* of instruction in learning words for motion. We had a two-way prediction based on inconclusive results in prior work. We predicted that speakers—independent of language and motion type—would show better learning when instructed with gesture+speech than in speech-only as gesture provides a second way of encoding new words and may reduce the cognitive load in a complex task such as novel word learning (Goldin-Meadow & Alibali, 2013; Kita et al., 2017). Alternatively, we predicted that speakers across languages and motion types would not show an effect of modality in learning pseudowords. This prediction was based on earlier research that suggested that iconic gestures might interfere with the ability to attach meaning to newly learned words by augmenting the semantic load already imposed by the novel spoken input (Emerson et al., 2016; Kelly & Lee, 2012).

## 2. Methods

### 2.1. Participants

Participants included 173 adult speakers, with either Chinese ($n$ = 60, $M_{age}$ = 19.20 [$SD$ = 0.93], *range* = 18–21, 35 females), English ($n$ = 53, $M_{age}$ = 19.00 [$SD$ = 1.25], *range* = 18–22, 45 females), or Turkish ($n$ = 60, $M_{age}$ = 20.83 [$SD$ = 1.76], *range* = 18–25, 36 females) as their native language. Originally, data were collected from 60 English speakers, but 7 participants were excluded due to experimental error (i.e., issues with data recording). The Chinese, English, and Turkish data were collected in Jingzhou City Hubei Province (China), Atlanta (USA), and Nevşehir (Turkey), respectively. The participants in each language had some knowledge of another language, having taken language courses as part of their secondary education—with English speakers learning Spanish or French and Turkish and Chinese speakers learning English. However, none of the participants had conversational fluency in a second language or took additional second language courses in college. Thus, the participants in each language were comparable in terms of their minimal exposure to a second language. The speakers of each language were also comparable in education: all were attending college at the time of the study. The participants were compensated by either course credit or small monetary compensation for their participation in the study.

### 2.2. Stimuli

The stimuli consisted of 16 motion animations that depicted motion events and 32 associated instructional videos that described these events in speech with or without gesture.

*Motion animations:* The motion animations depicted the path and manner of a star-shaped character in relation to a stationary spherical object. The star-shaped character performed motion with either different types of manner while the path of motion remained constant (i.e., *manner condition*, 8 animations) or performed motion with different types of path while the manner of motion remained constant (i.e., *path condition*, 8 animations). The animations used in this study were selected from a larger set of animations originally developed by Emerson et al. (2016), using Strata Design 3DCX6 software.

*Instructional videos:* The instructional videos consisted of 32 videos (16 for path, 16 for manner) of a male instructor, describing each animation with a pseudoword. Half of the instructional videos for each motion type described the motion in speech with gesture (*gesture + speech*; e.g., "frengu" + rapidly circling upward facing index finger in place to convey manner of rotating) and the other half without gesture (*speech-only*; "frengu"). The pseudowords that labeled the animations were originally developed by Emerson et al. (2016) and consisted of 8 nonsense words (*bripu, chulsu, derlu, frengu, lorpu, mernu, norcu*, and *sermu).* All pseudowords were articulated in a manner consistent with the phonetic pattern of each language by a native speaker. Furthermore, all pseudowords were disyllabic and were comparable to each other in terms of the number of phonemes as well as the type of surrounding phonological neighborhood based on PSIMETRICA (Mueller et al., 2003). We used the same 8 words to describe both manner and path variations, separately in the speech-only and gesture+speech conditions, to ensure that word labels did not influence speakers' learning (see Appendix A). All pseudowords were presented alone without any sentential context to avoid providing language-specific cues to the participants.[1]

### 2.3. Data collection

The experiment was conducted on a computer in a laboratory, and each participant was tested in their native language by a native speaker. At the beginning of the study, the participants were randomly assigned to one of the 4 between-subjects learning conditions: manner with speech-only ($n = 44$), (2) path with speech-only ($n = 44$), (3) manner with gesture+speech ($n = 45$), and (4) path with gesture+speech ($n = 40$). The participants in the speech-only condition learned pseudowords with speech-only instruction, while the ones in the gesture+speech condition learned them with gesture +speech instruction. The between-subjects design allowed us to assess learning within each category of motion type and modality type independently; it also

[1]As part of the original study, Emerson et al. (2016) tested *a* separate group of adult participants ($n$ = 57, $M_{age}$ = 20.43, *range* = 18–39, 15 males) to rate all pairs of manner ($n = 29$) and path words used ($n = 28$) on *a* 7-point Likert similarity scale (i.e., from 'not at all similar' to 'identical'). The similarity ratings for manner ($M$ = 3.28, $SD$ = 0.71) versus path ($M$ = 3.03, $SD$ = 0.74) animation pairs revealed no reliable differences ($t$ (55) = 1.30, $p$ = .198), showing that participants did not view the stimuli of one motion type (path or manner) as being more difficult to distinguish (i.e., more similar) than the other ($t$ (55) = 1.30, $p$ = .198; see Emerson et al., 2016 for more detail).

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

minimized potential practice and fatigue effects in learning as each participant only had to complete test items involving a single motion type within a single modality type.

Each participant completed 4 repeated blocks of learning with the same 8 pseudowords. The pseudowords labeled path variations in the *path condition* (either with speech or with gesture+speech) and manner variations in the *manner condition* (either with speech or with gesture+speech). The order of the pseudowords was randomized across participants and across the 4 blocks. The participant watched each of the 8 motion animations followed by the associated instructional video per block, one at a time. Halfway into each block, a mini-test on one of the words taught was administered to ensure that the participants were paying attention to the task at hand. At the end of each block, the participant was tested on 8 pseudowords using a forced-choice test. In the forced-choice test for each pseudoword, the participant was presented with two side-by-side animations accompanied by the instructor's voice that correctly labeled one of the animations. The participant was asked to choose the correct match by pressing a button on the computer keyboard; the associated buttons were marked with yellow tape (for choice of the animation on the left) and red tape (for choice of the animation on the right) for easy visibility by the participant. The placement of the correct animation on the right or the left of the computer screen as well as the presentation of the test trials for the pseudowords were randomized across participants. At each test trial, participants' accuracy rates (i.e., number of correctly chosen animations as the referent for the pseudoword) and reaction times (i.e., how quickly they pressed the associated button) were recorded (see Figure 1).

### 2.4. Data analysis

All responses were captured via a computer-based program (i.e., E-prime) with a maximum possible accuracy score of 8 per learning block along with a reaction time score for each response in milliseconds. We analyzed differences using two sets of repeated measure ANOVAs with learning (i.e., testing block) as a within-subject factor

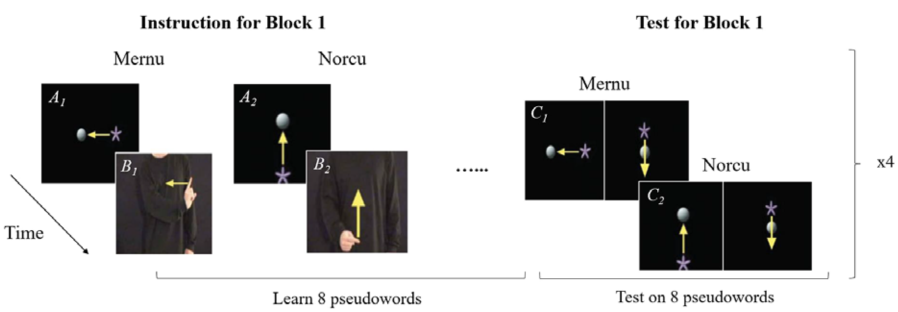



**Figure 1.** Sample experimental design showing learning blocks for the pseudowords mernu and norcu in the path with gesture + speech condition; ($A_1$) participant watches the motion animation for mernu; ($B_1$) participant watches the instructional video for mernu; ($A_2$) participant watches the motion animation for norcu; ($B_2$) participant watches the instructional video for norcu. This process is repeated for the remaining 6 pseudowords. ($C_1$) participant receives a block test for mernu while hearing the word mernu, ($C_2$) participant receives a test for norcu while hearing the word norcu. This process is repeated for the remaining 6 pseudowords. After the completion of Block 1, the whole process is repeated 3 more times with the same eight pseudoword, resulting in four blocks of accuracy and reaction time responses.

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

and language (English, Chinese, Turkish), motion type (manner, path), and modality (speech, gesture+speech) as between-subjects factors, separately for accuracy and reaction time of response. In a few of the blocks, the normality assumption was violated. However, in all of these cases, the skewness and kurtosis values remained within the acceptable range of normality (i.e., between −2 and + 2; George & Mallery, 2010) and the associated histograms showed a bell-shaped normal distribution, thus rendering ANOVA as the appropriate statistical tool for the analysis. All follow-up pairwise comparisons were adjusted, using Bonferroni correction.

## 3. Results

### 3.1. How accurately do speakers of different languages learn pseudowords for motion?

We first examined *accuracy rates* (i.e., correct response in each forced-choice test trial) in learning labels for motion events. As can be seen in Figure 2, accuracy improved with each block showing a main effect of learning ($F$ (3, 483) = 33.52, $p$ < .001, $\eta_p^2$ = 0.17). Speakers also differed in learning the labels across the three languages, showing a main effect of language ($F$ (2, 161) = 21.34, $p$ < .001, $\eta_p^2$ = 0.21). Chinese speakers showed lower accuracy across all four blocks of tests than both English and Turkish speakers (Bonferroni, $p$'s < .001), while the latter two did not differ in their accuracy rates (Bonferroni, $p$ = .79).

Accuracy rates also showed a main effect of motion type (i.e., manner versus path; $F$(1, 161) = 5.13, $p$ = .025, $\eta_p^2$ = 0.03), which did not interact with language ($F$(2, 161) = .71, $p$ = .50): Overall, speakers across all three languages showed slightly better learning of labels for manner than for path ($M_{manner}$ = 5.79, $SD$ = 1.67 versus $M_{path}$ = 5.38, $SD$ = 1.79).

On the other hand, accuracy rates showed no main effect of modality ($F$(1, 161) = .53, $p$ = .47) nor a Modality *x* Language interaction ($F$(2, 161) = .05, $p$ = .95). Speakers across the three languages showed comparable rates of learning when instructed with speech-only or with gesture+speech ($M_{speech}$ = 5.55, $SD$ = 1.73 versus $M_{gesture\ +\ speech}$ = 5.64, $SD$ = 1.75). We found no other two-, three- or four-way interactions between motion type, modality, language, and learning (see Table 1 for a full summary of statistical results for accuracy rates).

### 3.2. How quickly do speakers learn pseudowords for motion?

We next examined *reaction time* (i.e., response time in forced-choice test trials) in learning labels for motion events. As can be seen in Figure 3, response time decreased

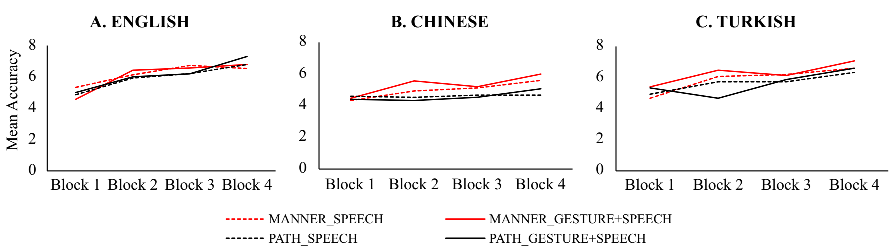



**Figure 2.** Mean accuracy scores of English (A), Chinese (B), and Turkish (C) speakers in each block of learning words for manner with gesture + speech (solid red lines), manner with speech-only (dotted red lines), path with gesture + speech (solid black lines) and path with speech-only (dotted black lines).

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

**Table 1.** Summary table for statistics on accuracy responses (significant effects are bolded)

| Variable | Sum of Squares | df | Mean Square | F | p |
|---|---|---|---|---|---|
| **Learning** | **186.012** | **3** | **62.004** | **33.52** | **<.001*****  |
| **Language** | **188.539** | **2** | **94.270** | **21.341** | **<.001***** |
| Modality | 2.333 | 1 | 2.333 | .528 | .468 |
| **MotionType** | **22.647** | **1** | **22.647** | **5.127** | **.025*** |
| Learning*Language | 22.439 | 6 | 3.740 | 2.02 | .06 |
| Learning*Modality | 4.282 | 3 | 1.427 | .772 | .51 |
| Learning*MotionType | 14.532 | 3 | 4.844 | 2.619 | .053 |
| Language * Modality | .455 | 2 | .227 | .051 | .950 |
| Language * MotionType | 6.246 | 2 | 3.123 | .707 | .495 |
| Modality * MotionType | 1.405 | 1 | 1.405 | .318 | .574 |
| Learning*Language*Modality | 8.480 | 6 | 1.413 | .764 | .598 |
| Learning*Language*MotionType | 10.366 | 6 | 1.728 | .934 | .47 |
| Learning*Modality*MotionType | 6.861 | 3 | 2.287 | 1.237 | .296 |
| Language * Modality * MotionType | 4.240 | 2 | 2.120 | .480 | .620 |
| Learning*Language*Modality*MotionType | 2.626 | 6 | .438 | .237 | .964 |

*p <.05; ***p <.001.

(i.e., got faster) with each block showing a main effect of learning ($F$ (2.19, 352.12) = 45.19, $p < .001$, $\eta_p^2 = 0.22$). Speakers differed in their reaction time in the three languages, showing a main effect of language ($F$ (2, 161) = 32.61, $p < .001$, $\eta_p^2 =$ 0.29). Chinese and Turkish speakers responded slower across all four blocks of tests than English speakers (Bonferroni, *p's* < .001).

Different from accuracy rates, reaction time showed no main effect of motion type (i.e., manner versus path; $F(1, 161) = .23$, $p = .63$) nor interaction between motion type and language ($F(2, 161) = .80$, $p = .45$). Speakers across the three languages showed similar reaction times when learning novel labels for manner or path of motion ($M_{manner} = 3512.02$, $SD = 1498.86$ versus $M_{path} = 3649.26$, $SD = 1668.38$). At the same time, motion type interacted with block ($F$ (2.19, 352.12) = 3.35, $p = .03$, $\eta_p^2 =$ 0.02): speakers across the three languages showed quicker response times in later blocks when learning words for manner than for path across all languages.

Similar to accuracy rates, reaction time showed no main effect of modality ($F$ (1, 161) = .812, $p = .37$) and no interaction between modality and language ($F$ (2, 161) = .212, $p = .81$). Speakers across the three languages responded at a similar

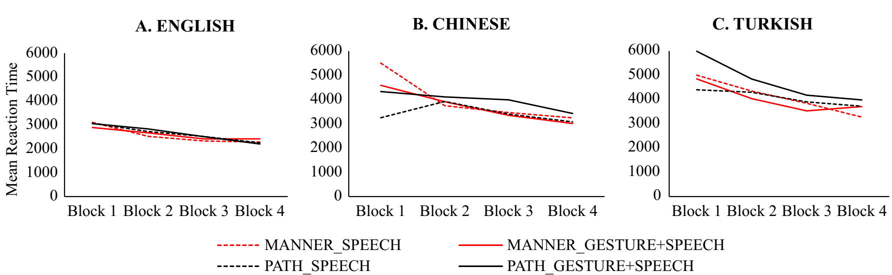



**Figure 3.** Mean reaction time scores of English (A), Chinese (B) and Turkish (C) speakers in each block of learning words for manner with gesture + speech (solid red lines), manner with speech-only (dotted red lines), path with gesture + speech (solid black lines) and path with speech-only (dotted black lines).

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

**Table 2.** Summary table for statistics on reaction time responses (significant effects are bolded)

| Variable | Sum of Squares | df | Mean Square | F | p |
|---|---|---|---|---|---|
| **Learning** | **122438925.955** | **3** | **55983522.138** | **45.18** | **<.001*****|
| **Language** | **310098669.521** | **2** | **155049334.760** | **32.605** | **<.001*****|
| Modality | 3862856.462 | 1 | 3862856.462 | .812 | .369 |
| MotionType | 1107875.725 | 1 | 1107875.725 | .233 | .630 |
| Learning*Language | 8326228.930 | 6 | 1387704.821 | 1.536 | .186 |
| Learning*Modality | 441191.364 | 6 | 147063.788 | .163 | .868 |
| **Learning*MotionType** | **9085501.614** | **3** | **3028500.538** | **3.353** | **.03*** |
| Language * Modality | 2016165.583 | 2 | 1008082.791 | .212 | .809 |
| Language * MotionType | 7631792.859 | 2 | 3815896.429 | .802 | .450 |
| Modality * MotionType | 11487285.690 | 1 | 11487285.690 | 2.416 | .122 |
| Learning*Language*Modality | 4833891.952 | 6 | 805648.659 | .892 | .476 |
| **Learning*Language*MotionType** | **14897491.058** | **6** | **2482915.176** | **2.749** | **.02*** |
| **Learning*Modality*MotionType** | **9570289.244** | **3** | **3190096.415** | **3.532** | **.02*** |
| Language * Modality * MotionType | 6067839.887 | 2 | 3033919.944 | .638 | .530 |
| Learning*Language*Modality*MotionType | 4813130.125 | 6 | 802188.354 | .888 | .478 |

*p <.05; ***p <.001.

speed when instructed with speech-only or with gesture+speech ($M_{speech}$ = 3476.23, $SD$ = 1433.26 versus $M_{gesture\ +\ speech}$ = 3687.18, $SD$ = 1724.55).

Our analysis of reaction time showed no four-way interaction but two 3-way interactions: The first was a Learning x Motion Type x Language interaction ($F(4.37, 352.12) = 2.75, p = .03, \eta_p^2 = 0.02$), which indicated that English speakers showed faster reaction times when learning manner or path verbs, compared to both Turkish and Chinese speakers; but this pattern was evident in all four blocks for manner verbs (Bonferroni, $ps \leq .02$) and only in the last three blocks for path verbs (Bonferroni, $ps \leq .01$). The second three-way interaction was between Learning x Motion Type x Modality ($F(2.19, 352.12) = 3.53, p = .03, \eta_p^2 = 0.02$), which showed that speakers in all three languages showed faster reaction times when learning path verbs in the first block but only in the speech-only condition (Bonferroni, $p = .02$; see Table 2 for a full summary of statistical results on reaction times).

## 3. Discussion

The world's languages follow a tertiary split in their expression of manner and path of motion—with greater expression of manner in S-languages (e.g., English), path in V-languages (e.g., Turkish), and comparable expression of manner and path in E-languages (e.g., Chinese; Slobin, 2004; Talmy, 2000). The cross-linguistic variability in motion descriptions has an effect on the nonverbal representation of motion, but this effect is evident during verbalization of the event in one's native language but is not present when *not* verbalizing the event (i.e., thinking-for-speaking account; Slobin, 1996). In this study, we asked whether the effect of language would or would not extend beyond verbalization of the motion event when learning pseudowords for motion by speakers of structurally different languages, particularly when the words were taught both with gesture and speech. More specifically, we asked whether learning pseudowords for motion would be affected by motion type (manner, path), language (Chinese, English, Turkish) and modality (speech-only, gesture+speech), using a word-learning paradigm that did not involve verbalization of the motion event in one's native language. Our results showed that speakers of all three languages learned pseudowords for manner and path—but with overall lower accuracy and slower response times in Chinese speakers. Regardless of the language they speak, participants learned pseudowords for manner more accurately than pseudowords for path, showing an effect of motion type. Their learning of words remained consistent, however, when instructed with gesture+speech or with speech only, thus showing no effect of modality of instruction.

### 3.3. Effect of language on learning novel words for motion

The speakers of all three languages learned the novel words for motion: they showed higher accuracy and faster reaction time in matching pseudowords to motion event animations over time. At the same time, Chinese speakers showed lower accuracy compared to both English and Turkish speakers and longer reaction times compared to English speakers in learning words for motion. What might explain the language effect that becomes evident in both measures of learning? One possible explanation might be the lexicalization of motion events in Chinese. Unlike English or Turkish speakers who rely on single verbs to express either manner or path of motion, Chinese speakers typically use serial verbs to express manner and path jointly (Paul, Emerson & Özçalışkan, 2022). In fact, as shown in earlier work, the majority of motion descriptions (62-86% across studies) by adult Chinese speakers relies on serial verb constructs encoding both manner and path in the verb (Chen & Guo, 2009; Tütüncü et al., 2023). The pseudowords in our study were all single words labeling either manner or path of motion—a pattern quite different than the habitual form of motion expression in Chinese. Accordingly, the single word labels might have resulted in lower word-like associations for Chinese speakers than they did for English or Turkish speakers. In fact, earlier research (Bartolotti & Marian, 2014) suggests that participants recognize and produce unword-like pseudowords less accurately and at slower speed than word-like pseudowords. The nature of the pseudowords in our study thus might have placed an extra cognitive processing load for Chinese speakers, resulting in lower accuracy and extended processing time in learning.

Another likely explanation is that the pseudowords used in our word-learning experiment were based on the Latin alphabet (e.g., mernu, norcu). Alphabetic languages such as English and Turkish are phonological, differing from Chinese, which is primarily ideographic. Ideographism in Chinese is largely conveyed by the special graphic quality of the Chinese characters (Gu, 2012). Previous research has shown that Chinese learners show slower reading times for English because they rely more on graphic and less on phonological cues compared to readers of an alphabetic language (Zhou, 1988). The greater reliance on ideographic cues, in turn, might have resulted in lower accuracy and reaction times in learning pseudowords by Chinese speakers in our study.

### 3.4. Effect of motion type in learning novel words for motion

We started with a two-way prediction for the effect of motion type (manner, path)—with the possibility of either an effect of language on learning manner versus path verbs or the lack of such an effect as would be predicted by the 'thinking for speaking account' (Slobin, 1996). The speakers in our study showed no differences in their learning of pseudowords by language, suggesting that the effect of language on learning does not go beyond verbalization—giving further support to Slobin's (1996) thinking for speaking account. More specifically, when instructed with novel words without any accompanying native speech production, S-, V- and E-language speakers in our study were able to learn words for both manner and path equally well. This finding is in line with previous findings that showed a lack of language effects on nonverbal representation of events when not verbalizing across a broad variety of nonverbal tasks (Cardini, 2010; Özçalışkan et al., 2016a, 2018; Papafragou et al., 2008; Skordos et al., 2020; Tütüncü et al., 2023). It also extended this work to the domain of word learning across structurally different languages. But why is there no effect of language on learning novel words for motion? As put forth by Chen and Guo (2009), speakers of all languages—regardless of structural differences—have the lexical means to encode both manner and path components of a motion event. As aforementioned, we see evidence of this in silent gestures (i.e., gestures produced without speech), where speakers of different languages show cross-linguistic similarities in their expression of manner and path of motion by encoding both in gesture; Özçalışkan et al., 2016a, 2018, 2024b; Tütüncü et al., 2023), suggesting that these two event components are available to speakers for encoding motion across different languages. And when given a task with no time constraints and no verbalization of the event, they were able to learn labels for either motion component at comparable rates.

At the same time, even though the speakers in our study showed similar patterns in learning words for manner and path in each language, they also showed better learning for words encoding manner than path of motion *across* languages. What might explain the better performance in learning pseudowords for manner, particularly given Talmy's (2000) proposal that path constitutes a core component of a motion event in event construal. According to Talmy, path information must always be explicitly encoded—regardless of whether it is expressed in the main verb or in a particle associated with the verb—whereas the overt expression of manner is optional. If that is the case, one could expect speakers of all languages to learn pseudowords for path better than the ones for manner. And indeed, some of the earlier work provided some evidence for the path bias, with speakers of either language type showing better learning or recall of path than manner information (Emerson et al., 2016; Gennari et al., 2002; Maguire et al., 2010; Skordos et al., 2020). At the same time, however, more recent work (Aktan-Erciyes et al., 2022) suggests that the salience of an event component might also play an important role in learning. Aktan-Erciyes and colleagues found that Turkish speakers (V-language) rated manner information to be more salient than path information in animated motion scenes that depicted both components of motion, even if they found the verbal expression of path easier. The same participants also showed an effect of manner—but not path—when asked to make similarity judgements (without verbalization). More specifically, the Turkish speaker's similarity judgements were affected by differences in manner salience but not path salience, with differences in manner but *not* path resulting in decreased perceived similarity. These findings thus suggest that not only typological structure (i.e., path constituting the core component of a motion event), but also salience in the depiction of a motion component could jointly influence speakers' perception and learning of different motion components of an event.

In addition, the depiction of path in our animations included not only encoding the moving entity (star-shaped figure) but also a landmark (e.g., a spherical entity) in relation to which the figure moved. As such, the learning of words for path required not only paying attention to the moving entity but also the object that served as the goal or source for the motion—thus differing from manner animations, which necessitated focusing only on the moving figure. The added cognitive load of paying attention to two versus one entity might have resulted in a lower rate of learning in the path condition than in the manner condition. However, this difference was not evident in reaction times, suggesting equal processing times for both manner and path components of a motion event in the minds of all speakers independent of language. The lack of congruence between the two measures of learning highlights the need for future research that can include other conditions (e.g., animations with both manner and path with or without landmarks) to shed further light on the source of the differences we observed in our study.

Another reason for the difference could be the design of our study. Different from earlier work, most of which used a within-subject design in manner versus path conditions (e.g., Gennari et al., 2002; Özer & Göksun, 2020; Papafragou et al., 2002), our study relied on a between-subjects approach in which each participant either learned words only for path or only for manner. This, in turn, might have allowed them to more easily focus exclusively on one of the two motion components (manner or path) during the word-learning task, resulting in better learning for manner words.

### 3.5. Effect of modality in learning novel words for motion

Gesture is an integral aspect of communication and provides an important window into the mind. It has now been shown that gestures can facilitate learning especially if the task at hand is complex (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Perry & Elder, 1997). Learning a new language can be complex for beginners, and gesture, in turn, may facilitate language learning in two ways: it may ease the cognitive load by presenting a second modality (i.e., visual presentation) that is different than speech alone (i.e., auditory presentation); it may also allow the learners to explore ideas that may be difficult to comprehend or to verbalize with speech alone (Goldin-Meadow, 2000). However, contrary to this, our results showed no beneficial effect of gesture+speech instruction over speech instruction. One possible explanation for the absence of a gesture effect on learning could be the nature of the task. The participants were asked to learn eight pseudowords in total over four blocks, which may have resulted in a near-ceiling effect. Indeed, participants' performance on average was fairly high (76% correct) across trials and languages (English: 83%, Chinese: 64%, Turkish: 81%) even in the speech-only condition, leaving gestures relatively little room to improve the performance already achieved by the instruction of the words in speech alone. Future studies that test gesture's role in learning a greater number of words that are also more complex can tell us more about the contribution of task complexity on the beneficial effect of gesture in learning novel words.

Another explanation for the lack of a gesture effect comes from children learning new words in their native language. There is considerable work that suggests close coupling between child gesture use and subsequent word learning, in which early gestures (mostly points at objects) precede and predict the time of onset and the size of children's early spoken vocabularies (e.g., Iverson & Goldin-Meadow, 2005; Özçalışkan et al., 2017). Importantly however, most of this earlier work on word learning focused on unique gestures (i.e., referents conveyed only in gesture but not yet in speech), but not on gesture+speech combinations. There is in fact work that suggests that the size of children's vocabularies at 42 months is predicted by their unique gesture vocabularies at 14 months but *not* by their gesture+speech combinations at 18 months (Rowe & Goldin-Meadow, 2009; see also Rowe et al., 2008). These findings thus suggest that the beneficial effect of gesture in learning words might be less evident if the instruction involved a gesture+speech combination (as in our study)—a possibility that needs to be tested in future studies.

In addition, in our study, the participants learned pseudowords, all ending with a similar sound (i.e., all ending with the phoneme /u/). As such, gesture+speech combinations might have made it more difficult to attach meaning to these highly similar phonetic forms, thus eliminating the possible enhancing effects that gesture can provide. In other words, the additional semantic content provided by the iconic gestures may have interfered with the ability to attach meaning to the newly learned pseudowords. Since speech is phonetically novel in our task, additional meaning provided by iconic gestures may have taxed the learners' cognitive system. Indeed, previous research found that when learning phonetically hard pairs of words, gesture instructions do not help learners; in fact, they hinder their performance (Kelly & Lee, 2012). This explanation also fits well with the second language learning model (Baddeley et al., 1998) which proposes that the phonological loop in working memory is dedicated to learning a new language; and when the phonological loop is taxed with novel speech sounds, this disrupts the encoding of those novel sounds into permanent memories for new words. Kelly and Lee (2012) argue that, considering the already taxed load of encoding in the working memory, addition of iconic gestures may have added a visually distracting dimension to the task of learning. Although this does not compromise the ability to learn these sounds and later remember them, it may nonetheless have eliminated the boost that iconic gestures could have brought to learning.

There is also research that suggests that gesture aids learning, especially when the learners produce the gestures rather than just observe them—a pattern that has been shown both for adult (Engelkamp & Dehn, 2000; Macedonia et al., 2011; Morett, 2014) and child (Cook & Goldin-Meadow, 2006; Goldin-Meadow et al., 2012; Tellier, 2008) learners. Similar to this earlier work, in our study, participants only viewed the gestures without producing them themselves. Future research that systematically vary the joint effect of both observing and producing gesture can further illuminate the relative contribution of gesture to learning novel words for different types of instructional gesture exposure.

In conclusion, our study extended previous work showing a lack of a language-specific effect on nonverbal representation of events when not verbalizing the event in one's native language (Cardini, 2010; Özçalışkan et al., 2016a, 2018; Papafragou et al., 2008; Skordos et al., 2020; Tütüncü et al., 2023) to the domain of word learning across structurally different languages. Even though Chinese speakers performed less accurately than English and Turkish speakers and slower than English speakers, their learning of manner and path pseudowords did not interact with language. This study also took the word-learning paradigm one step further by also examining the effect of instruction type, showing no advantage of gesture+speech instruction over speech-only instruction in learning. These findings thus highlight that simply

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

observing gesture with novel words might not be sufficient to facilitate learning novel words for motion beyond the instruction provided with speech alone.

**Data availability statement.** https://osf.io/x3wb6/?view_only=fa3b1ebd577743e58510409d60ccc6d5.

**Acknowledgements.** English data collection was supported by an GSU RCALL doctoral fellowship to Emerson. Turkish data collection was supported by a grant from The Scientific and Technological Research Council of Turkey (TUBITAK; 53325897-115.02-170549) to Şengül.

**Competing interest.** The authors declare none.

## References

Aktan-Erciyes, A., Akbuğa, E., Dik, F. N., & Göksun, T. (2022). Linguistic and nonlinguistic evaluation of motion events in a path-focused language. *Applied Psycholinguistics*, 43(4), 829–865.

Alibali, M. W., & Goldin-Meadow, S. (1993). Modeling learning using evidence from speech and gesture. In *Proceedings of the fifteenth annual conference of the cognitive science society* (pp. 203–208). Erlbaum.

Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition*, 102, 16–48.

Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, 37(2), 286–309.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158–173.

Bartolotti, J., & Marian, V. (2014). Wordlikeness and novel word learning. In *36th proceedings of the annual meeting of the cognitive science society*. The Cognitive Science Society.

Cardini, F. E. (2010). Evidence against Whorfian effects in motion conceptualisation. *Journal of Pragmatics*, 42(5), 1442–1459.

Chen, L., & Guo, J. (2009). Motion events in Chinese novels: Evidence for an equipollently-framed language. *Journal of Pragmatics*, 41(9), 1749–1766.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, 23(1), 43–71.

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development*, 7(2), 211–232.

De Knop, S., & Dirven, R. (2008). Motion and location events in German, French and English: A typological, contrastive and pedagogical approach. *Applications of Cognitive Linguistics*, 9, 295.

De Knop, S., & Gallez, F. (2011). Manner of motion: A privileged dimension of German expressions. *International Journal of Cognitive Linguistics*, 2(1), 25–40.

Emerson, S. N., Conway, C. M., & Özçalışkan, Ş. (2020). Semantic P600—but not N400—effects index crosslinguistic variability in speakers' expectancies for expression of motion. *Neuropsychologia*, 149, 107638.

Emerson, S. N., Limia, V. D., & Özçalışkan, Ş. (2021). Cross-linguistic transfer in Turkish–English bilinguals' descriptions of motion events. *Lingua*, 264, 103153.

Emerson, S. N., Özçalışkan, Ş., & Frishkoff, G. A. (2016). Effects of motion type and modality on word learning in English. *Applied Psycholinguistics*, 37(3), 643–671.

Engelkamp, J., & Dehn, D. M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 671.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. (2002). Motion events in language and cognition. *Cognition*, 83(1), 49–79.

George, D. & Mallery, M. (2010). *SPSS for windows step by step: A simple guide and reference, 17.0 update (10th ed.)*. Pearson.

Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development*, 71(1), 231–239.

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, 64, 257–283.

Goldin-Meadow, S., Levine, S. C., Zinchenko, E., Yip, T. K., Hemani, N., & Factor, L. (2012). ng gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science*, 15(6), 876–884.

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105(27), 9163–9168.

Gu, M. D. (2012). *Sinologism: An alternative to orientalism and postcolonialism*. Routledge.

Gullberg, M., Hendriks, H., & Hickmann, M. (2008). Learning to talk and gesture about motion in French. *First Language*, 28(2), 200–236.

Hickmann, M., Taranne, P., & Bonnet, P. (2009). Motion in first language acquisition: Manner and Path in French and English child language. *Journal of Child Language*, 36(4), 705–741.

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience*, 19(7), 1175–1192.

Huang, X., Kim, N., & Christianson, K. (2019). Gesture and vocabulary learning in a second language. *Language Learning*, 69(1), 177–197.

Hupp, J. M., & Gingras, M. C. (2016). The role of gesture meaningfulness in word learning. *Gesture*, 15(3), 340–356.

Ibarretxe-Antunano, I. (2009). Path salience in motion events. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, & S. Özçalışkan (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 403–414). Psychology Press.

Ibarretxe-Antuñano, I. (2012). Linguistic typology in motion events: Path and manner. *Anuario del Seminario de Filología Vasca 'Julio de Urquijo'. International Journal of Basque Linguistics and Philology*, 41(3), 841–849.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16(5), 367–371.

Kang, S., & Tversky, B. (2016). From hands to minds: Gestures promote understanding. *Cognitive Research: Principles and Implications*, 1(1), 1–15.

Kelly, S. D., & Lee, A. L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*, 27(6), 793–807.

Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General*, 139, 638–653. https://doi.org/10.1037/a0020507.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, 124(3), 245–266.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48(1), 16–32.

Kushch, O., Igualada, A., & Prieto, P. (2018). Prominence in speech and gesture favour second language novel word learning. *Language, Cognition and Neuroscience*, 33(8), 992–1004.

Lewandowski, W., & Özçalışkan, Ş. (2018). How event perspective influences speech and co-speech gestures about motion. *Journal of Pragmatics*, 128, 22–29.

Lewandowski, W. & Özçalışkan, Ş. (2023). Running across the mind or across the park: does speech about physical and metaphorical motion go hand in hand? *Cognitive Linguistics*, 34(3-4), 411–444. https://doi.org/10.1515/cog-2022-0077

Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping*, 32(6), 982–998.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Haryu, E., Vanegas, S., … Sanchez-Davis, B. (2010). A developmental shift from similar to language-specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. *Cognition*, 114(3), 299–319.

Mathias, B., Sureth, L., Hartwigsen, G., Macedonia, M., Mayer, K. M., & von Kriegstein, K. (2021). Visual sensory cortices causally contribute to auditory word recognition following sensorimotor-enriched vocabulary training. *Cerebral Cortex*, 31(1), 513–528.


https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

Morett, L. M. (2014). When hands speak louder than words: The role of gesture in the communication, encoding, and recall of words in a novel second language. *The Modern Language Journal*, 98(3), 834–853.

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1353.

Naigles, L. R., & Terrazas, P. (1998). Motion-verb generalizations in English and Spanish: Influences of language and syntax. *Psychological Science*, 9(5), 363–369.

Oh, K. J. (2003). *Language, cognition, and development: Motion events in English and Korean*. University of California.

Özçalışkan, Ş. (2009). Learning to talk about spatial motion in language-specific ways. In J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, K. Nakamura, & Ş. Özçalışkan (Eds.), *Cross-linguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 263–276). Psychology Press.

Özçalışkan, Ş. (2015). Ways of crossing a spatial boundary in typologically distinct languages. *Applied Psycholinguistics*, 36(2), 485–508.

Özçalişkan, Ş. (2016). Do gestures follow speech in bilinguals' description of motion? *Bilingualism: Language and Cognition*, 19(3), 644–653.

Özçalışkan, Ş., Adamson, L. B., Dimitrova, N., & Baumann, S. (2017). Early gesture provides a helping hand to spoken vocabulary development for children with autism, Down syndrome and typical development. *Journal of Cognition and Development*, 18(3), 325–337.

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016a). Does language shape silent gesture? *Cognition*, 148, 10–18.

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016b). Is seeing gesture necessary to gesture like a native speaker? *Psychological Science*, 27(5), 737–747.

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2018). Blind speakers show language-specific patterns in co-speech gesture but not silent gesture. *Cognitive Science*, 42(3), 1001–1014.

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2024a). What the development of gesture with and without speech can tell us about the effect of language on thought. *Language & Cognition*, 16(1), 220–241. https://doi.org/10.1017/langcog.2023.34

Özçalışkan, Ş., Lucero, C. & Goldin-Meadow, S. (2024b). Is vision necessary for the timely acquisition of language-specific patterns in co-speech gesture and their lack in silent gesture? *Developmental Science*, 27(5), pp. e13507. https://doi.org/10.1111/desc.13507

Özçalışkan, Ş. & Slobin, D. (2003). Codability effects on the expression of manner of motion in Turkish and English. In A. S. Özsoy, D. Akar, M. Nakipoglu-Demiralp, E. Erguvanli-Taylan, & A. Aksu-Koç (Eds.), *Studies in Turkish linguistics* (pp. 259–270). Bogaziçi University Press.

Özçalışkan, Ş., & Slobin, D. I. (1999). Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In *Proceedings of the 23rd annual Boston University conference on language development* (Vol. 2, pp. 541–552). Cascadilla Press.

Özer, D., & Göksun, T. (2020). Visual-spatial and verbal abilities differentially affect processing of gestural vs. spoken expressions. *Language, Cognition and Neuroscience*, 35(7), 896–914.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184.

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle,'n'roll: The representation of motion in language and cognition. *Cognition*, 84(2), 189–219.

Papafragou, A., & Selimis, S. (2009). On the acquisition of motion verbs crosslinguistically. In M. Baltazani, G. K. Giannakis, T. Tsangalidis & G. J. Xydopoulos (Eds.), *Proceedings of the 8th international conference on Greek linguistics* (pp. 351–365). Ioannina: University of Ioannina.

Paul, J., Emerson, S. N., & Özçalışkan, Ş. (2022). Does dialect matter in the expression of motion? Ways of encoding manner and path in Babao and standard Mandarin. *Lingua*, 270, 1, 103215–17.

Perry, M., & Elder, A. D. (1997). Knowledge in transition: Adults' developing understanding of a principle of physical causality. *Cognitive Development*, 12(1), 131–157.

Ping, R., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology*, 44(5), 1277–1287.

Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, 12(1), 182–187.

https://doi.org/10.1017/langcog.2024.44 Published online by Cambridge University Press

Rowe, M. L., Özçalışkan, Ş., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, 28(2), 182–199.

Shafto, C. L., Havasi, C., & Snedeker, J. (2014). On the plasticity of semantic generalizations: Children and adults modify their verb lexicalization biases in response to changing input. *Developmental Psychology*, 50 (3), 794–808.

Skordos, D., Bunger, A., Richards, C., Selimis, S., Trueswell, J., & Papafragou, A. (2020). Motion verbs and memory for motion events. *Cognitive Neuropsychology*, 37(5-6), 254–270.

Skordos, D., & Papafragou, A. (2014). Lexical, syntactic, and semantic-geometric factors in the acquisition of motion predicates. *Developmental Psychology*, 50(7), 1985–1998.

Slobin, D. (1996). From 'thought' and 'language' to 'thinking for speaking. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press.

Slobin, D. I., (2004). The many ways to search for a frog: linguistic typology and the expression of motion events. In S. Strömqvist, & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 219–257). Lawrence Erlbaum Associates.

Smyder, R., & Harrigan, K. (2021). The influence of language-specific and universal factors on acquisition of motion verbs. *Proceedings of the Linguistic Society of America*, 6(1), 927–937.

Sweller, N., Shinooka-Phelan, A., & Austin, E. (2020). The effects of observing and producing gestures on Japanese word learning. *Acta psychologica*, 207, 103079.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and semantic description. Vol. III: Grammatical categories and the lexicon* (pp. 36–149). Cambridge University Press.

Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring*. The MIT Press.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, 8(2), 219–235.

Tusun, A., & Hendriks, H. (2019). Voluntary motion events in Uyghur: A typological perspective. *Lingua*, 226, 69–88.

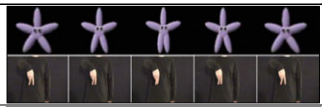
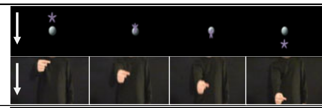
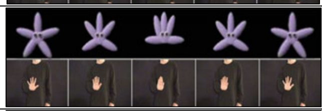
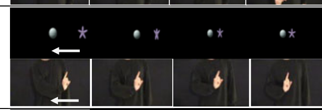
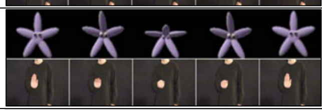
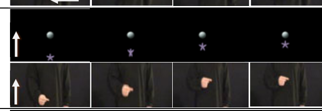
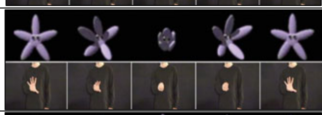
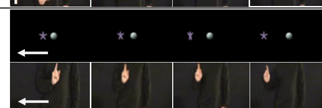
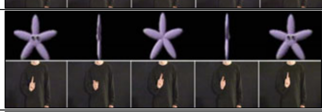
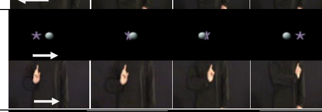
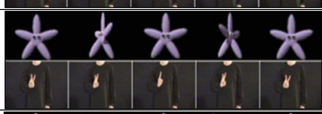
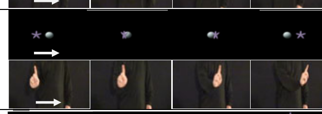
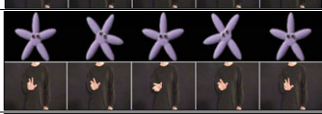
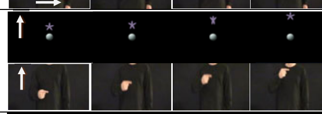
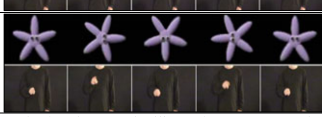
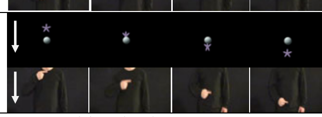
Tütüncü, I. S., Paul, J., Emerson, S. N., Şengül, M., Knezevic, M., & Özçalışkan, Ş. (2023). When gestures do or do not follow language-specific patterns of motion expression in speech: evidence from Chinese, English and Turkish. *Cognitive Science*, 47(4), e13261.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28(2), 187–204.

Zhou, M. (1988). "Script effects and reading strategies: Ideographic language readers vs. alphabetic language readers in ESL". Dissertations and Theses. *Portland State University. Department of English*, Paper 3950. https://doi.org/10.15760/etd.5834

## APPENDIX A

**Table A1.** Pseudowords with their related gesture instructions (Manner videos (A), Path videos (B); stimuli (1), gesture instruction (2))

| Pseudoword | | A. Manner | B. Path* |
|---|---|---|---|
| bripu | 1.<br>2. |  |  |
| chulsu | 1.<br>2. |  |  |
| derlu | 1.<br>2. |  |  |
| frengu | 1.<br>2. |  |  |
| lorpu | 1.<br>2. |  |  |
| mernu | 1.<br>2. |  |  |
| norcu | 1.<br>2. |  |  |
| sermu | 1.<br>2. |  |  |

*The arrows on gestural depictions of animations are for illustration purposes and were not part of the animation.

**Cite this article:** Tütüncü, I. S., Emerson, S. N., Paul, J., Şengül, M., & Özçalışkan, Ş. (2024). Learning novel words for motion by speakers of structurally different languages, *Language and Cognition*, 1–22. https://doi.org/10.1017/langcog.2024.44